International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 9 Issue 2, Mar-Apr 2025 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Presentation of Statistical Outputs from Counting Approach: Shall Frequency Come Before Percentage?

Nana Celestin

Foundation of Applied Statistics and Data Management (FASTDAM), Buea, Cameroon

ABSTRACT

The general approach to describe discrete / categorical variables is using counting techniques that are summarized into frequencies and proportions / percentages. When dealing with a conceptual component made of complementary indicators, Multiple-Responses Sets (MRS) as composite variable can be generated using Multiple-Responses Analysis (MRA) technique; but whether individual indicator's score or composite-aggregated scores within conceptual component, the statistical outputs are generated in term of frequency and proportion. But the statistical outputs for discrete / categorical variables are presented differently by different scholars, statisticians or researchers as if there was no set standard. Following convincing scientific and statistical analysis and demonstrations, it is proven that the acceptable standard is proportion coming before frequency [e.i. 50% (30) instead of 30 (50%)] given the mathematic conservation law that scientifically gives priority to proportion, which is termed proportionate-conservation law in this article. I therefore recommend that this standard be followed for the uniformity of the presentation of statistical outputs for discrete / categorical variables.

KEYWORDS: *Discrete / categorical variables; Analysis; Counting techniques; Outputs; Frequency; Proportion / Percentage; Standard*

How to cite this paper: Nana Celestin "Presentation of Statistical Outputs from Counting Approach: Shall Frequency Come Before Percentage?" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-9 | Issue-2, April 2025, pp.812-822, URL: www.ijtsrd.com/papers/ijtsrd78356.pdf

IJTSRD78356

Introduction and background

Statistics is present in research, educational, social, political and economic sectors of our society and constitutes an important tool for decision-making for States, individuals and organizations at local, national and international levels. Fair and updated statistics and database are indispensable if our vision for economic development, social welfare and equity should be achieved in prospect of sustainable development. We need precise statistical methods to gather and reduce the abstraction of data or information, to discuss the well-being of hypotheses and make predictions, to describe variability in a population or to infer whether differences among individuals, samples or populations are due to chance or not, and to estimate unknown-deductive constants, as well as to determine for a data set assertive limits (confidence levels, confidence intervals) within which the true value is more likely to fall. In fact, if research can be defined as a scientific exercise consisting of a systematic collection, analysis, interpretation and discussion of data to answer a certain question or solve a problem, it is therefore clear that the backbone of research is statistics and applied statisticians with their holistic comprehension and application of research methods and statistical analysis better fit in this context. Whether basic or fundamental research (necessary to generate new theories to address unresolved or potential problems) or applied / operational research (necessary to verify how efficiently already defined theories can help solving identified problems), statistics remains a key tool. It is always difficult to separate statistics from probability and other mathematical theories, given that most of statistical laws have their roots in probabilistic theories. But nowadays, we do not need to be highly vested with those mathematical theories behind statistical formulae to use statistical tools. This is because the majority of commonly used statistical procedures have been simplified to arithmetic calculations. In fact, the development of statistical software that follows the development in electronic devices and computers has been of great contribution to the development of research and such software are always characterized by the patent simplification of long and complex mathematical theories to simple statistical calculations. However, it is important to be

aware that accurate and sufficient knowledge of statistical theories and principles are always essential in the use of statistical software. These facilities and simplifications in the application and use of statistical tools are of great advantage to research, because experts of various fields, without necessarily being professionals of statistical data analysis, have now developed themselves and have some familiarities with basic statistics, which they apply to their profession, and seek for the advice of professional statisticians only when really necessary, and can properly work with them in a participatory way if they so do. This enables them to collaborate better with professional statisticians who may not have sufficient idea of their fields. Statistical analysis is one of the key aspects of research from the beginning to the end. At the beginning, the research plan takes into account the type and the size of data to be analyzed and the statistical tool that will be used to analyze them. At the end, data will be analyzed, the statistical outputs interpreted, commented and discussed. Statistics is a set of scientific methods which are used to collect, organize, analyze, discuss, summarize and present data. Some of these methods can enable us to draw conclusions and take decisions. It applies both to fundamental and applied (operational) research. Scientific because it abides to rigorous or rigidly accurate set rules or standards that follow a logical and step-by-step sequences to achieve a given objective. The scientific nature of statistics responds to the methodological-technical assumption / requirements, which means it contributes to yield results that are valid and verifiable because they are based on evidence and follow rules and procedures, that make them verifiable and critical. It is often asserted that statistics should be empirical, thus in line with the determinism principle, which means that results are based on evidence or grounded from real life experience or observations; this is what is generally expected, though some predictive or simulative studies may rely on imaginative / fictitious data within the framework of a preconceived hypothetical model, based primarily on surmise or incomplete evidence rather than real life findings. The statistical approach and the overall research design are specific to study context or social situation, whereby the ingredients put together to serve the research process and achieving the set objectives shall form a balanced system tailored to fit the study context, thus responding to the ontological assumption; that is why sense of logical, systematic and conceptual reasoning as well as sense of specificity should be the major characteristics of an applied statistician or statistical applications in general. But the essential is not to yield statistical outputs whether qualitative or quantitative, but to further reduce the abstraction of information through appropriate presentation of results or findings, as to enhance their comprehension and exploitation by the general public.

Statement of the Problem

The general approach to describe discrete / categorical variables is using counting techniques that are summarized into frequencies and proportions. When dealing with a conceptual component made of complementary indicators, MRS as composite variable can be generated using MRA technique. But whether individual indicator's scores or composite-aggregated scores within conceptual component, the statistical outputs are presented in term of frequency and proportion. But the statistical outputs for discrete / categorical variables are presented differently by different scholars, statisticians or researchers as if there was no set standard. This article aims at resolving this dilemma using convincing scientific and mathematic demonstrations.

Research question

What standard approach shall be used to present statistical outputs, generally frequency and proportion generated by the analysis of discrete / categorical variables?

Conjecture

Frequency shall be place before proportion, for instance 20 (50%).

Methodology

The approach used in this article is purely analytical and demonstrative.

Demonstrations

This consists essentially of conceptually and critically presenting facts in a systematic sequential manners to uplift equivoques on a given scientific questions or approve or reject it. Using rigorous analytical, mathematical and experimental approaches, you can accept or reject research questions, conjectures, assumptions or hypotheses.

What is being measured?

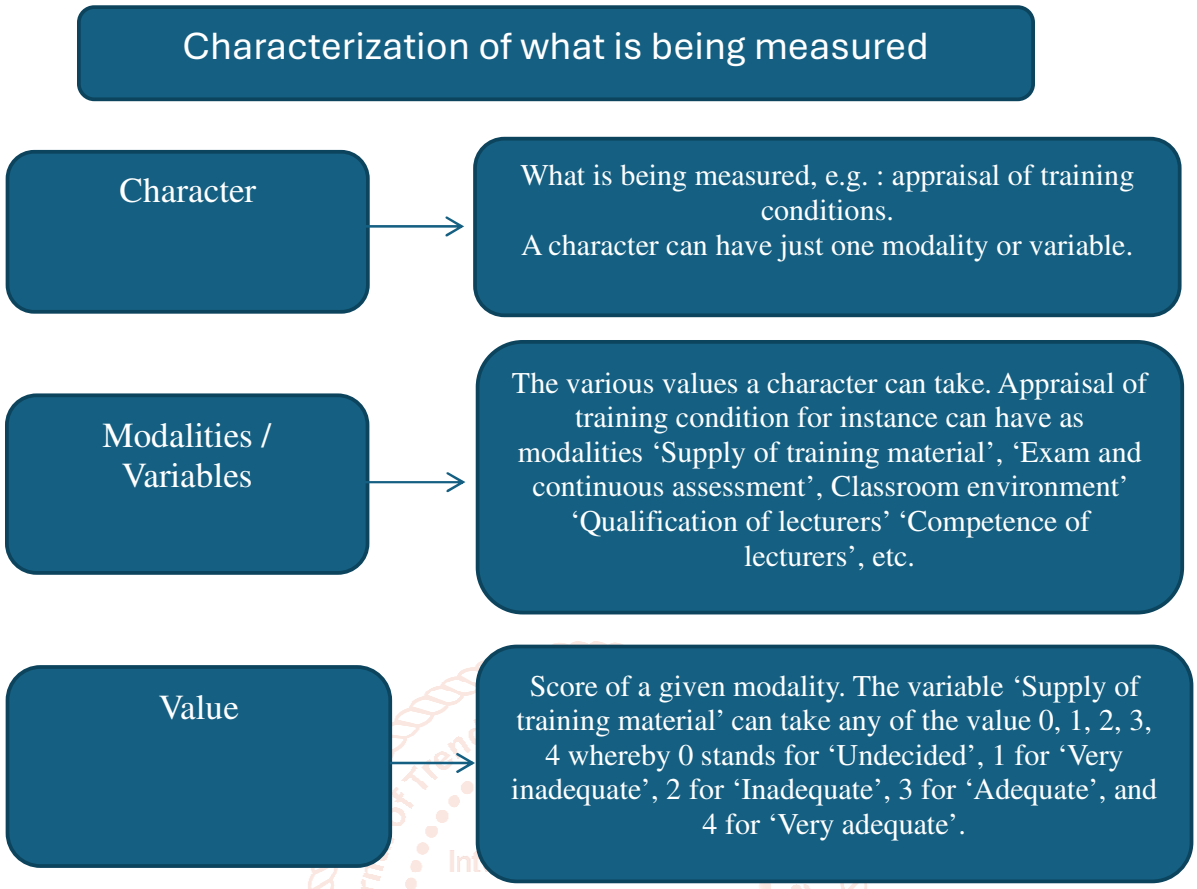

Figure 1: What are we measuring?

Measurement levels of variables

There are three measurement levels of variables as presented on figure 2.

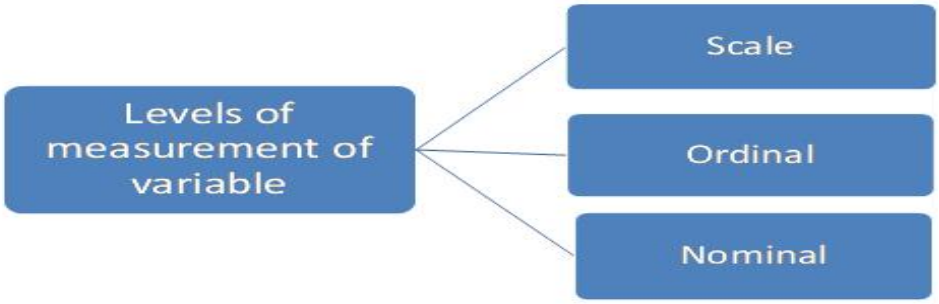

Figure 2: Measurement levels of variables

We are essentially concerned with categorical variables, whether ordinal or nominal in this article. Table 1 below further describes the measurement levels of variables such that we can make a clear-cut difference between scale (continuous variable) and discrete / categorical variable which can be ordinal or nominal.

Table 1: Measurement levels, equivalent terms and main characteristics

Level	Equivalent terms	Main characteristics
Nominal	Categorical, qualitative, discrete	Categories have no set rank order. For instance, if we attribute the categories 1, 2 and 3 to respectively Ancestority / Kamitism, Christianity and Islam, the codes 1, 2 and 3 here just help to identify the types of religion and are not numerically meaningful because they are not quantitatively comparable. Such coding scheme most often based on dummy numerical values is used just to facilitate data collection, management and analysis.

Ordinal	Discrete, categorical	Categories can be rank-ordered from low to high. E.g.: the variable 'Age' on a scale or continuous measurement level takes the following possible values in years (rounded up to 1 decimal), 1.0, 2.0, 3.5, 5.5, 5.5, 6, 6, 7.5, 7.3, 8.0, 24.0, 25.0, 35.0, 70.0, 72.0…'Age' now can be organized into age groups or age ranges as follows: 1 '1-10', 2 '11-20' 3 '21-30' 4 '31-40' 5 '41-50' 6 '51-60' …. Age ranges are categories of the variables 'Age' and we can say that 'Age' has been transformed from scale to categorical variable that we can name in our data base 'AgeRg' to differentiate it from the scale variable. 'AgeRg' is categorical but ordinal because the categories can be rank-ordered from the smallest to the highest in a numerically meaningful manner. In fact, a fellow who falls in the category 1 is deliberately younger or less aged as compared to somebody who falls in the category 3. N.B. What makes the difference between age group and age range is that age group deals with standard ranges within a given field while age ranges just refer to contextualized categories defined to meet some statistical requirements. In the medical field for instance, 0-59 months is a standard age group. Those aged above 36 years can no longer be employed in the public service. For this reason, using 36 years as cut point in a study context considering this parameter can lead to age groups because a standard is followed.
Scale	Numerical, interval, ratio, continuous, quantitative	Gradable values are expressed in numerical terms and are organized in order—from the lowest to the highest value or vice versa on an interval or ratio scale. Ratio scale takes only positive values while interval scale takes both positive and negative values. Coefficient of variation is not computed with interval-scale variables.

Description of variables / descriptive statistics

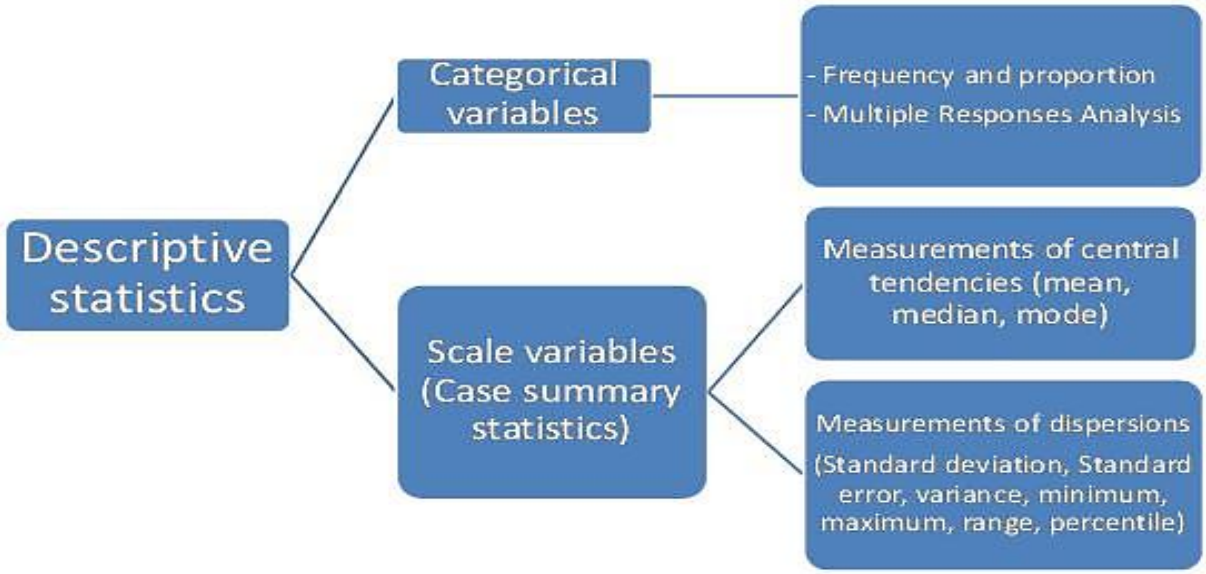

Figure 3: Descriptive statistics

Descriptives for categorical variables: Frequency and proportion

Table 2: General presentation of a frequency table

X_i	n_i	f(i) %
0	3	15
1	4	20
2	2	10
3	5	25
4	6	30
Total	20	100

Frequencies:

This consists of counting the number of occurrences in a contingency table.

The various symbols used are:

N: entire sample or population;

X_i: variable's value for the i^{th} row of the table;

With i = 1, 2, 3, 4,……, m where m is the possible number of categories.

n=The number of count within a given category.

f(i)=Proportion or percentage for the i^{th} category or row of the table.

Formula to calculate proportion:

P (%)=(n / N) * 100

Where n is the frequency or the number of count and N the total sample.

For the category X_3, n_3=5

P_{X3} =(5/20)*100

=25%

Where:

n=5

N=20

Table 3: Example of frequency table with gender as the variable

Gender	Frequency	Percent	Valid Percent
Female	15	24.6	24.6
Male	46	75.4	75.4
Total	61	100.0	100.0

The tabulation along with frequency and percentage gives a useful description of data sets for variables with ordered or unordered categories.

Gender is the label of the variable. The label of the variable should be properly and explicitly written, because it is what refers to the variable in most statistical outputs.

The first column contains the values or categories of the variable (sex).

The frequency column indicates the number of people or cases in each category notably female and male.

The percentage column represents the percentage or proportion taken by each category. These percentages are based on the total sample size and the frequency or number of cases within each category. The valid percentage column contains the percentages calculated only with those who really gave valid responses and of course excludes missing values. Valid percentage is what is normally used. Valid percent is not necessary in the case of this table because there is no missing value. Cumulative percentage is valid only when dealing with ordinal variables and it gives the aggregated percentage for a set of category levels and is not necessary for this table handling a nominal variable.

Categorical variables unlike scale variables are described using essentially frequency and proportion though composite variable generated from then can be scale / continuous. However, composite variable or aggregated scores using Multiple-Responses Set (MRS) is freely recommended because though more complex to compute, it is statistically proven to be more accurate with 100% precision. The composite or aggregated scoring is generally done considering adjustment for conceptual polarization whereby negative are reversed into positive or vice-versa for uniformity and consistency in conceptual polarity. Below is the formula for computing MRS. It is therefore recommended that if faced with triangulation is measurement levels in a study whereby aggregate score from a conceptual component is unavoidable, scale variables shall instead be transformed or recoded into categorical variables for cross-tabulations, though aggregating by summing to generate a composite scale / continuous variable from a conceptual component mathematically can be tolerated.

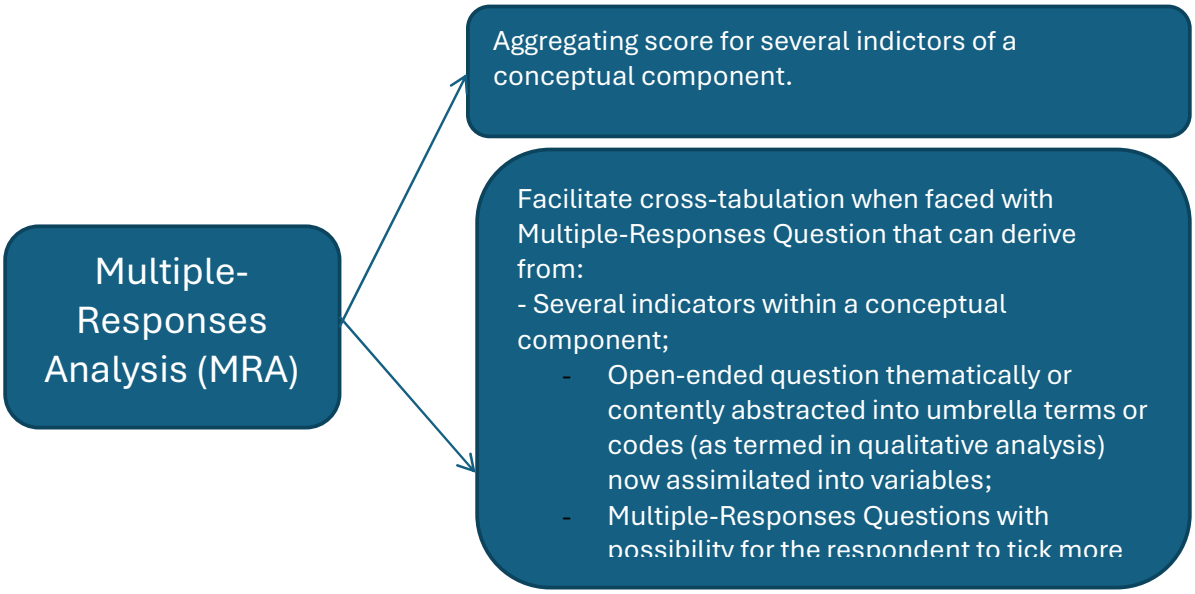

Figure 4: Multiple-Responses Analysis

Table 4: Formula steps to calculate MRS

\Conceptual Component A		L1	L2	L3	N
X_1		$n_{L1}\ X_1$	$n_{L2}\ X_1$	$n_{L3}\ X_1$	$N\ X_1 = n_{L1}\ X_1 + n_{L2}\ X_1 + n_{L3}\ X_1$
X_2					$N\ X_2$
X_3					$N\ X_3$
X_4					$N\ X_4$
X_5					$N\ X_5$
X_6					$N\ X_6$
X_7		$n_{L1}\ X_7$	$n_{L2}\ X_7$	$n_{L3}\ X_7$	$N\ X_7$
Aggregate (MRS)	$n_{responses}$	$n_{L1}\ X_1 + \ldots + n_{L1}\ X_7$	$n_{L2}\ X_1 + \ldots + n_{L2}\ X_7$	$n_{L3}\ X_1 + \ldots + n_{L3}\ X_7$	$N_{responses} = \Sigma\ (n_{L1}\ X_1 + \ldots + n_{L1}\ X_7) + (n_{L2}\ X_1 + \ldots + n_{L2}\ X_7) + (n_{L3}\ X_1 + \ldots + n_{L3}\ X_7)$
	%	$(n_{responses}L1 / N_{responses})*100.$	$(n_{responses}L2 / N_{responses})*100.$	$(n_{responses}L3 / N_{responses})*100.$	% L1 + %L2 + %L3=100.0 %

Table 5: Perceived importance of statistics by FASTDAM's students and lecturers

Perception of statistics	Agree	Disagree	N
Develop the sense of logical and systematic reasoning	97.4% (487)	2.6% (13)	500
Is important in research	88.6% (443)	11.4% (57)	500
Helps to understand the social world	82.2% (411)	17.8% (89)	500
Helps to plan development	71.4% (357)	28.6% (143)	500
Reduces existential abstraction and subjectivity	64.6% (323)	35.4% (177)	500
It is difficult*	17.8% (89)	82.2% (411)	500
Not really necessary*	19.6% (98)	80.4% (402)	500
Aggregated score (MRS) without reversing conceptual polarity (This is wrong)	63.1% (2208)	36.9% (1292)	3500
Aggregated score (MRS) with reversed conceptual polarization (*)	81.0% (2834)	19.0% (666)	3500

How do we now present output in prose writing?
Output from table 3 for instance can be presented in two different ways:

Frequency before proportion approach (first approach):
Males were more than the females with proportion of 46 (75.4%) as against 15 (24.6%) for the female.

Proportion before frequency approach (second approach):
Males were more than the females with proportion of 75.4% (46) as against 24.6% (15) for the female.

Which approach is recommendable?
This is where the proportionate-conservation law gives the edge to the second approach.

Practical demonstration:
Let us consider a sample of 60 males and 40 females. To the question 'do you like statistics?' 30 males answered 'yes' and 20 female answered 'yes'. If we considered frequency, the tendency is to resolve that males like statistics more females. Calculating their respective proportion gives us 50% for males and 50% for female thus implying that males and females have equal level of interest for statistics. Proportion here clearly appears to portray the context more precisely than frequency, thus shall be given priority in data presentation, thus abiding to the proportionate-conservation law. This freely rejects the conjecture earlier stated.

Another application of the proportionate-conservation law: Chi-Square test of independence
This test is appropriate when one would like to test the independence of two variables. For instance, FASTDAM would like to compare preferences of male and female FASTDAM's students for three different employers. In the survey, a sample of male and female students from FASTDAM had to be chosen between government, private sector and self-employment. The findings are presented on the table below.

	Government	Private	Self-employed	Total
Male	50.0% (30)	33.3% (20)	16.7% (10)	60
Female	66.7% (40)	25.0% (15)	8.3% (5)	60
Total	55.0% (55)	30.0% (30)	15.0% (15)	100

Below is the structure of the contingency table. In fact, a contingency table is generally a R x C table (rows times columns). This table is a 2 x 3 contingency table as we have 2 rows and 3 columns for a degree of freedom (d.f.) of 2 (2-1 * 3-1).

	Government	Private	Self-employed	Total
Male	C1,1	C1,2	C1,3	R1
Female	C2,1	C2,2	C2,3	R2
Total	C1	C2	C3	T

This is the Chi-Square formula that we will apply to determine whether preference of employers is independent of gender or not.

$$X^2 = \sum \frac{(O-E)^2}{E}$$

O=Observed Frequency
E=Expected Frequency

	Government	Private	Self-employed	Total
Male	50.0% (30)	33.3% (20)	16.7% (10)	60
Female	66.7% (40)	25.0% (15)	8.3% (5)	60
Total	55.0% (70)	30.0% (35)	15.0% (15)	120

	Government	Private	Self-employed	Total
Male	E1,1	E1,2	E1,3	R1
Female	E2,1	E2,2	E2,3	R2
Total	C1	C2	C3	T

We have to compute the expected frequencies for each bloc or cell using the following formula:

$$\text{Expected value} - \frac{Row\ sum\ x\ column\ sum}{grand\ total}$$

$$E_{1,1} = \frac{60 \times 70}{120} = 35 \quad E_{1,2} = \frac{60 \times 35}{120} = 17.5 \quad E_{1,3} = \frac{60 \times 15}{120} = 7.5$$

$$E_{2,1} = \frac{60 \times 70}{120} = 35 \quad E_{2,2} = \frac{60 \times 35}{120} = 17.5 \quad E_{2,3} = \frac{60 \times 15}{120} = 7.5$$

Proportion is conservative and explains the rationale behind computing expected values using the formula above. For $C_{1,1}$ for instance, a total of 55 students out of 100 preferred government as employer and since they are 60 female students, one would expect if the null hypothesis was true (60x70)/120 or 35 male students to prefer government as employer and the same for the female.

	Government	Private	Self-employed	Total
Male	35%(30)	17.5% (20)	7.5% (10)	60
Female	35% (40)	17.5% (15)	7.5% (5)	60
Total	70	35	15	120

We now have to move to the next stage of our computation, and to make things easy for ourselves, we place the expected values calculated besides their corresponding observed values as presented on the table. Using the formula presented above, let us calculate the test value.

$$\chi 2 = \frac{(30-35)^2}{35} + \frac{(20-17.5)^2}{17.5} + \frac{(10-7.5)^2}{7.5} + \frac{(40-35)^2}{35} + \frac{(15-17.5)^2}{17.5} + \frac{(5-7.5)^2}{7.5}$$

$$= 3.8096$$
$$= 3.810$$

The final step consists of calculating the degree of freedom and to position the test value on the chi-square distribution table and read the P-value and then take the decision whether to accept or reject the hypothesis. Chi-square test is generally a right-tailed test.

D.f. = (R-1)(C-1)
= (2-1)(3-1)
= 1*2
= 2

On the Chi-Square distribution table, the corresponding value for Alpha=0.05 at d.f.=2 is 5.991.

3.810 < 5.991 therefore accepting the null hypothesis that preference for employer is statistically not significantly dependent of gender.

The P-value corresponding to a Chi-Square test value of 3.810 is 0.149 > 0.05, therefore implying that male and female statistically have the same preference for employer.

Context and facts: Analyzing indicators and generating composite variable for subjective indicators in a liker-scale design

When the predictor is a component made of several indicators, the standard statistical approach recommends that we appreciate the statistical trend indicator by indicator and generate a composite variable that represents the aggregated score for the conceptual component.

Example:

How would you rate the training conditions you experienced? *Please rate each of the applicable options and place a tick under the appropriate level.*

very bad						very good	
1	2	3	4	5	6	7	
☐	☐	☐	☐	☐	☐	☐	Design of the training programme
☐	☐	☐	☐	☐	☐	☐	Assistance/advice for your final examination
☐	☐	☐	☐	☐	☐	☐	Exams, continuous assessment (overall grading system)
☐	☐	☐	☐	☐	☐	☐	Training quality of lecturers
☐	☐	☐	☐	☐	☐	☐	Equipment and stocking of library
☐	☐	☐	☐	☐	☐	☐	Supply of training material
☐	☐	☐	☐	☐	☐	☐	Accommodation and catering facilities

The predictor "Training Condition" is a conceptual component made of 7 indicators.

We will appreciate the statistics for the seven indicators taken individually and then generate a composite variable that aggregates the score of the seven indicators using different approaches.

Theoretical-weighted-mean approach applied to individual indicators
Using the simple proportion, estimate the score of each respondent for the indicator on a scale of 7.

Sum the score of all the respondents for the indicator and divide it by the total number of respondents to obtain the average score. Some scholars, depending on the study context and the set operational objectives can decide to set the cut point higher, above average.

Knowing that the theoretical average score for an indicator on a scale of 7 is 3.5, you can now distribute your group score for the indicator into two groups. The group of those whose scores is below average and the group of those whose score is equal to or above average, which average is set as cut point. The second group can be roughly classified as satisfied and the first group as dissatisfied. If the first respondent ticked 1 and the second respondent 5, then the sum of their scores is 6 and the average 3, below the average cut-point, indicating that the group perception falls under unsatisfactory.

Theoretical-weighted-sum approach applied to individual indicators
Compute the sum of score for all the cases / respondents and obtain the value A. For 2 respondents, you expect a maximum score of 14.

Using simple proportion, weight A on scale of 100 knowing that the maximum score for the 2 respondents is 14. If the first respondent ticked 1 and the second respondent 5, then the sum of their score is 6.

14 → 100

A → X

X will represent the score of these respondents on a scale of hundred.

Using the formula above, the weight on a scale of 100 is 42.9%, below 50% thus weighing toward unsatisfactory for that indicator.

Counting and proportion-weighted approach applied to individual indicators
If the first respondents ticked 1 and the second respondent 5, after collapsing, it is clear that the first one falls under unsatisfactory and the second one under satisfactory making up 50% satisfactory and 50% unsatisfactory which at first sight is more real and freely perceptible, thus contradicting the theoretical-weighted-sum or theoretical-weighted-mean approach applied by some scholars.

Theoretical-weighted-mean approach applied to the entire cohort and component
This approach consists in summing the scores of all the respondents, calculating the average and verifying if the value is above the theoretical average, and assessing the deviation from the mean using standard deviation.

With this approach, if we have the following responses from the two respondents for the seven indicators:
Respondent A has scored 3, 3, 3, 4, 7, 7, 7 for a total score of 34
Respondent B has scored 2, 2, 5, 6, 6, 6, 7 for a total score of 34

Mathematically, these two respondents have the same appreciation of the situation but a policy maker looking at the situation with a more pragmatic and realistic perception may not see it that way.

To him, respondent B seems to be more comfortable or satisfied than respondent A because he is actually satisfied with 5 indicators out of 7 as against three for respondent A.

In fact, for the two respondents, we would have expected a maximum of (7 * 7) + (7 * 7) =98 points for an average of 49. Summing the score of the two respondents, we have 68 for an average 34 which falls below the average cut point of 49 points indicating that the situation is perceived unsatisfactory.

Contextual-weighted-mean approach applied to the entire cohort and component
In this approach, respondents' scores are classified as falling above or below the group average. In the case above, the contextual-weighted-mean is 34. The two respondents will be classified as average and the situation perceived as fair or fairly good. The Contextual mean is equivalent to class average. Categorizing using the contextual-weighted-mean seems to slightly correct the error or bias from the theoretical-weighted-mean approach which clearly classifies the situation as unsatisfactory.

Counting and proportion-weighted approach applied to the entire cohort and the conceptual component using MRA (Multiple-Response Analysis)

We can collapse using the recode command the seven-level scales into 3 main categories as follows:

1-3=Unsatisfactory

4= Fairly satisfactory

5-7= Satisfactory.

With the very responses:

Respondent A has scored 3, 3, 3, 4, 7, 7, 7 for a total score of 34.

Respondent B has scored 2, 2, 5, 6, 6, 6, 7 for a total score of 34

1-3=Unsatisfactory = 5 for a response-weight of (5/14)*100= 35.7%.

4= Fairly satisfactory = 1 for a response-weight of (1/14)*100= 7.1%.

5-7= Satisfactory = 8 for a response-weight of (8/14)*100= 57.1%.

Indicators' scores and Multiple Responses Set (aggregated indicators' scores) are tabulated as followed:

Indicators		1-3	4	5-7	Total responses
Design of the training program		1	1	0	2
Assistance/advice for your final examination		2	0	0	2
Exams, continuous assessment (overall grading system)		0	0	2	2
Training quality of lecturers		0	0	2	2
Equipment and stocking of library		1	0	1	2
Supply of training material		0	0	2	2
Accommodation and catering facilities		1	0	1	2
Aggregate (MRS)	n	5	1	8	14
	%	35.7%	7.1%	57.1%	100.0%

We have a weight of 57.1 % for the category 5-7 (satisfactory) which indicates that in general the situation is satisfactory thus contradicting the theoretical -weighted-mean approach and to a smaller extent the contextual-weighted-mean approach applied to the entire cohort as explained above.

In social research, given the fact that we are dealing with non-tangible indicators, counting approach is highly recommended as compared to the mean-weighted methods in generating composite variables. Using MRA, we can estimate with 100% precision the weight of each of the category levels (1...7). We can crosstab MRA result (MRS) with independent variables like sex and use Chi-Square test to compare proportions for significant difference with the support of Epi Info version 6.04d or any other appropriate software application sometimes termed calculators.

MRA freely corrects the bias of theoretical-weighted-mean approach applied to the entire cohort as explained above because it proceeds essentially by simple and multiple counting supported by the proportionate-conservation law. This tool and its applications are quite essential in socio-psychological research mostly when faced with non-tangible indicators on a Likert-scale. Moreover, people can freely make a tangible and more realistic sense out of the statement '57.1% of the responses fall under satisfied' than 'the scores of people falls above the average cut point'. Of course, MRS is more time consuming in computation and summary of outputs in statistical tables; but it remains the most appropriate or reliable approach. Also, the theoretical-weighted-mean approach uses the Standard Deviation as an estimate of the variability of sample's perceptions or deviation from average. This in a psychometric context is incomplete because it does not tell us whether the deviation weighs more toward positive or negative view. The Counting Approach associated to MRA corrects this bias because for instance, we can freely appreciate the weights of those that strongly agree or strongly disagree. For instance, if two indicators have the same proportion when collapsing agree and strongly agree, the one having the higher proportion for strongly agree deliberately is more emphasized upon and the one having the higher proportion of strongly disagree the lesser emphasized upon, thus freely interpreted as conceptually skewing toward either positive or negative views or perceptions.

Analyzing ranking data

MRA and its statistical conservative property still amplifies the process and corrects for bias when aggregating score in the context of individual-indicator-scoring approach which aggregated process is termed frequency-weight approach of course as earlier explained, more reliable than the sum-weight approach. In this approach, instead of defining combination-categories (permutation, combination, tree diagram), simply provide a field to enter the rank score for each of the options to be ranked. For instance if there are 4 options to be ranked, define each of the option as a variable that can take any of the possible ranking code from 1 to 4, 4 been the highest. Simply analyze the data using frequency and proportion approaches and distribute the outputs in a two-ways-variable table, one way standing for the options that were ranked and the other way for the rank levels. Using MRA, you can aggregate score within rank levels as to have the ranking for the entire component while within and across indicators, the proportions freely determine the trend.

Conclusion and recommendation

Following convincing scientific and statistical analysis and demonstrations, it is proven that the acceptable standard in presenting statistical outputs for the description of discrete / categorical variables is proportion coming before frequency given the mathematic conservation law that scientifically gives priority to proportion, which is termed proportionate-conservation law in this article. Uniformity in scientific presentation is recommendable, reason why applying the most scientifically convincing approach as for proportionate-conservation law in this paper, whereby the most conservative shall come first, which is proportion in the context of this demonstration, as it conservatively has an edge over frequency. In fact, as proven by this analysis and demonstration, frequency can be silent in favor of proportion in prose wring without affecting the real interpretative context, whereas silencing proportion leaving frequency can lead to erroneous or biased interpretation.

References

[1] Bluman, A. G. (2001). *Elementary statistics: A step-by-step approach (fourth edition)*. New York: Mc Graw Hill.

[2] IBM SPSS Statistics. *Statistical Package for Social Sciences (SPSS)*, version 20 upwards.

[3] Nana, C. (2018). *Research Methods and Applied Statistics: beginners and Advanced Learners* (3rd Edition). Buea: GOOAHEAD.